8 Referências Bibliográficas

1 Baeza-Yates, R; Ribeiro-Neto, B; **Modern Information Retrieval**. Addison-Wesley-Longman, 1999;

2 Bayes, T; Price, R; 1763; **An Essay Towards Solving a Problem in The Doctrine of Chance**; Philosophical Transaction of The Royal Society of London, 53, pp 370-418.

3 Bergholz, A; Chidlovskii, B; 2003; **Crawling the Hidden Web**; Proceedings of the Fourth Internation Conference on Web Information Systems Engineering (WISE 2003).

4 Bergman, M; **The Deep Web: Surfacing Hidden Value and Semantization**; Research Challenges in Information Science, Pags 225-236, 22-24 April 2009.

5 Bhattachary, I; Geetor, L; Licamele, L; 2006; **Query-time Entity Resolution (poster)**; In Conference on Knowledge Discovery and Data Mining (KDD), pp 529-534, Philadelphia PA, 2006.

6 Bilenko, M; Mooney, R; **Adaptative Duplicate Detection Using Learnable String Similarity Measures**; Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2003), Washington DC, pp 39-48, August, 2003.

7 Bilenko, M**; Learnable Similarity Functions and Their Applications to Record Linkage and Clustering**; Ph.D. proposal, Department of Computer Sciences, University of Texas at Austin, October 2003, 47 pages.

8 Bleiholder, J; Naumann, F; **Data Fusion**; ACM Computing Surveys, Vol. 41, No. 1, Article 1, Publication date: December 2008.

9 Brauner, D; Gazola, A; Casanova, M; Breitman, K; 2007; **An Instance-based Approach for Matching Export Schemas of Geographical Database Web Services**; IX Brazilian Symposium on GeoInformatics (GeoInfo 2007), Brazil.

10 Brauner, D; Gazola, A; Casanova, M; Breitman, K; 2008; **Adaptative Matching of Database Web Services Export Schemas**; 10th International Conference on Enterprise Information Systems (ICEIS 2008), 12-16, June, 2008, Barcelona, Spain

11 Carvalho, M; Laender, A; 2009; **Abordagens Evolucionárias para Problemas Relacionados à Integração de Dados**; Belo Horizonte: UFMG , 2009 (Tese de Doutorado)

12 Cohen W; Ravikumar, P; Fienberg, S; 2003; **A Comparison of String Distance Metrics for Name-Matching Tasks**; Proceedings of IJCAI-03 Workshop on Information Integration, pp 73-78.

13 Cover, T; Hart, P; 1967; **Nearest Neighbor Pattern Classification**; IEEE Transactions on Information Theory, vol. 13, Issue 1, pp 21-27, Jan 1967.

14 Davis, C; Salles, E; 2007; **Approximate String Matching for Geographic Names and Personal Names**; IX Brazilian Symposium on GeoInformatics, Campos do Jordão, Brazil, November 25-28, 2007, INPE p. 49-60.

15 Dice, L; 1945; **Measures of the Amount of Ecologic Association Between Species**; Ecology, Vol. 26, No. 3(Jul., 1945), pp. 297-302.

16 Dong, X; Halevy, A; Madhavan, J; 2005; **Reference Conciliation in Complex Information Spaces**; In Conference on the Management of Data (SIGMOD), pp 85-96, Baltimore, MD, 2005.

17 He, B; Patel, M; Zhang, Z; Chang, K; 2007; **Accessing The Deep Web: A Survey**; Communications of the ACM, Vol. 50, Issue 5, Pags 94-101, May 2007

18 Hernandéz, M; Stolfo, S; 1995; **The Merge/Purge Problem for Large Databases**; In Conference on the Management of Data (SIGMOD), pp 127-138, San Jose, CA, 1995.

19 Jaccard, P; 1901; **Étude Comparative de La Distribution Florale Dans Une Portion dês Alpes ET dês Jura**; Bulletin de La Socieété Vaudoise dês Sciences Naturelles 37, PP 547-579.

20 Jaro, M; 1989; **Advances in Record Linkage Methodology as Applied to the 1985 Census of Tampa Florida**; Journal of The American Statistical Society, pp. 84

21 Köpke, H; Rahm, E; 2009**; Frameworks for entity matching: A comparison**; Data Knol, Eng 2009, doi:10.1016/j.datak.2009.10.003.

22 Krause, E; 1987; **Taxicab Geometry**; Dover. ISBN 0-486-25202-7

23 Laender, A; Silva, A; Ribeiro-Neto, B; Teixeira, J; 2009; **A Brief Survey of Web Data Extraction Tools**; Belo Horizonte: UFMG, 2009.

24 Leme, L; Casanova, M; 2009; **Conceptual Schema Matching Based on Similarity Heuristics**; Rio de Janeiro: PUC-Rio, 2009 (Tese de Doutorado).

25 Levenshtein, V; 1966; **Binary Codes Capable of Correcting Deletions, Insertions and Reversals**. Soviet Physics Doklady 10, 8, 707-710.

26 Manning, C; Raghavam, P; Schütze, H; **Introduction to Information Retrieval**; Cambridge University Press, 2008, chapter 1

27 Merril, D; 2009; Mashups: **The new breed of Web app**; http://www.ibm.com/developerworks/xml/library/x-mashups.html; acessado em 24/07/2009.

28 Monge, A; Elkan, C; 1996; **The Field Matching Problem: Algorithm and Applications**; In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining.

29 Newcombe, H; Kennedy, J; Axford, S; James, A; **Automatic Linkage of Vital Record**; Science, Vol. 130 (16 October 1959), pp. 954-959.

30 Ntoulas, A; Zerfos, P; Cho, J; 2005; Downloading Hidden Web Content; UCLA Computer Science, 2005.

31 Oguri, P; 2006 ; **Machine Learning For Sentiment Classification**; Dissertação de Mestrado, Departamento de Informática, Puc-Rio, 2006

32 Peisu, X; Ke, T; Qinzhen, H; 2008; **A Framework of Deep Web Crawler**; Proceedings of the 27th Chinese Control Conference, July 16-18, 2008, Kunming, Yunnan, China.

33 Pereira, D; Ribeiro-Neto, B; 2009; **A Web-Based Framework for Entity Resolution And Creation of Authority Files**; Belo Horizonte: UFMG , 2009 (Tese de Doutorado).

34 Salton, G; Wong, A; Yang, C; 1975; **A Vector Space Model for Automatic Indexing**; Communications of the ACM, vol. 18, n. 11, pp 613-620

35 Spärk Jones, K; 1972; **A Statistical Interpretation of Term Specificity and Its Applications in Retrieval**; Journal of Documentation 28, V1, pp. 11-21.

36 Su, W; Wang, J; Lochovsky, F; 2010; **Record Matching over Query Results from Multiple Web Databases**; IEEE Transactions on Knowledge and Data Engineering, Vol 22, No.4, April 2010.

37 Raghavan, S; Garcia-Molina, H; 2001; **Crawling the Hidden Web**; Proceeedings of the 27th International Conference on Very Large Data Bases (VLDB), pags 129-138, 2001.

38 Russel, R.; 1918; **United States Patent Office, N.1.261.167**

39 Wang, J; Wen, J; Lochovsky, F; Ma, W; 2004; **Instance-based Schema Matching for Web Databases by Domain-specific Query Probing**; The 30th Very Large Data Base Conference (VLDB'04).

40 Winkler, W; 1999; **The State of Record Linkage and Current Research Problems**; Statistics of Income Division, Internal Revenue Service Publication R99/04.

41 Wright, A; 2004; **In Search of the Deep Web**; Salon.com; http://www.salon.com/tech/feature/2004/03/09/deep_web/

42 Vapnik, V; 1998; **Statistical Learning Theory**; Wiley, 1998.

43 Zhang, H; 2004; **The Optimality of Naïve bayes**; In 17th International Florida Artificial Intelligence Research Society Conference, 2006.